arXiv:1004.3946v1 [math.NA] 22 Apr 2010

On efficiency of Orthogonal Matching Pursuit.[1]

Eugene Livshitz[2]

Abstract

We show that if a matrix Φ satisfies the RIP of order $[CK^{1.2}]$ with isometry constant $\delta = cK^{-0.2}$ and has coherence less than $1/(20K^{0.8})$, then Orthogonal Matching Pursuit (OMP) will recover K-sparse signal x from $y = \Phi x$ in at most $[CK^{1.2}]$ iterations. This result implies that K-sparse signal can be recovered via OMP by $M = O(K^{1.6} \log N)$ measurements.

1 Introduction.

The emerging theory of Compressed Sensing (CS) has provided a new framework for signal acquisition [1], [3], [8]. Let us recall some basic concepts of CS. Let Φ be a $M \times N$ matrix ($M < N$). The basic problem in CS is to construct a stable and fast algorithm for recovery a signal $x \in \mathbb{R}^d$ that has K non-zero components (K-sparse signal) from measurements $y = \Phi x \in \mathbb{R}^M$ and to determine (M, N, K) for which such algorithms exist.

E. Candés and T. Tao proved that Basic Pursuit (BP)

$$\widehat{x}(y) = \operatorname{argmin}\{|z|_1 : \ \Phi z = y\}.$$

can provide the exact recovery of arbitrary K-sparse $x \in \mathbb{R}^N$ by $M = O(K \log(N/K))$ measurements.

In this article we study signal recovery via Orthogonal Matching Pursuit(OMP). Although theoretical results for OMP are essentially worse than for BP, its computational simplicity allows OMP to achieve very good result in practise [18].

Algorithm: Orthogonal Matching Pursuit

Input: Φ, y.
Initialization: $r^0 := y$, $x^0 := 0$, $\Lambda^0 = \emptyset$, $l = 0$.
Iterations: Define

$$\Lambda^{l+1} := \Lambda^l \cup \operatorname{argmax}_i |\langle r^l, \phi_i \rangle|,$$

$$x^{l+1} := \operatorname{argmin}_{z:\ \operatorname{supp}(z) \subset \Lambda^{l+1}} \|y - \Phi z\| \quad r^{l+1} := y - \Phi x^{l+1}.$$

If $r_{l+1} = 0$, stop. Otherwise let $l := l + 1$ and begin a new iteration.
Output: If algorithm stops at l-th iteration, output is $\widehat{x} = x_l$.

By ϕ_i, $1 \le i \le N$, we denote the i-th column of Φ. We assume that $\|\phi_i\| = 1$, $1 \le i \le N$. To formulate results on recovery via OMP we use two basic properties of matrix Φ.

[1]This research is partially supported by Russian Foundation for Basic Research project 08-01-00799 and 09-01-12173

[2]E-mail: evgliv at gmail.com

- Coherence of Φ
$$\mu(\Phi) := \sup_{i\neq j} |\langle \phi_i, \phi_j\rangle|.$$

- Restricted Isometry Property ([4]). A matrix Φ satisfies Restricted Isometry Property (RIP) of order K with isometry constant $\delta \in (0,1)$ if the inequality
$$(1-\delta)\|x\|^2 \le \|\Phi x\|^2 \le (1+\delta)\|x\|^2$$
holds for all K-sparse $x \in \mathbb{R}^N$.

It's well known (see [11], [12]) that if
$$\mu(\Phi) < \frac{1}{2K-1}$$
then OMP will recover arbitrary K-sparse signal x from $y = \Phi x$ in exactly K iterartions. The stability of recovery via OMP in the term of coherence of Φ has been studied in [11], [18], [9], [10], [17], [15]. Recently M. Davenport and M. Wakin [6], and E. Liu and V.N. Temlyakov [14] showed that if Φ satisfies RIP of order $K+1$ with isometry constant
$$\delta = \frac{1}{3K^{1/2}}([6]), \quad \delta = \frac{1}{(1+2^{1/2})K^{1/2}}([14]),$$
then OMP recovers arbitrary K-sparse signal $x \in \mathbb{R}^N$ in exactly K iterations.

To compare these results we recall estimates on coherence and RIP for normalized random Bernoulli matrices Φ (each entry is $\pm M^{-1/2}$ with probability $1/2$). For rather big c_μ we have with high probability that
$$\mu(\Phi) \le c_\mu M^{-1/2} \log^{1/2} N. \tag{1}$$
R. Baraniuk, M. Davenport R. Devore and M. Wakin [2] (see also earlier B.S. Kashin's work [13]) showed that random Bernoulli matrix Φ with high probability satisfy RIP of order K with isometry constant δ for
$$M = O\left(\frac{K\log(N/K)}{\delta^2}\right). \tag{2}$$
Thus both results require $M = O(K^2)$ measurements for recovery of K-sparse signal. The aim of this article to show that OMP can recover sparse signals by essentially less number of measurements.

Theorem 1. *There exist absolute constants $C > 0$ and $c > 0$ such that if Φ satisfies the RIP of order $[CK^{1.2}]$ with isometry constant $\delta = cK^{-0.2}$ and has coherence $\mu(\Phi) \le 1/(20K^{0.8})$, then for any K-sparse $x \in \mathbb{R}^N$, OMP will recover x exactly from $y = \Phi x$ in at most $[CK^{1.2}]$ iterations.*

Inequalities (1) and (2) imply that for rather big absolute constant $C_M > 0$ with high probability normalized random Bernoulli matrix Φ with
$$M = \left[C_M K^{1.6} \log N\right]$$

satisfies condition of Theorem 1.

Much less is known about the lower estimates. H. Rauhut [16] proves that if $M \le \widetilde{c}K^{3/2}$ then for most random $M \times N$ matrices there exists a K-sparse signal $x \in \mathbb{R}^N$ that can not be recovered via K iterations of OMP. Moreover, it's conjectured in [16] (see also [5]) that for $M \le \widetilde{c}_n K^{2-1/n}$, $n \in \mathbb{N}$, with high probability there exists a K-sparse signal $x \in \mathbb{R}^N$ that can not be recovered via K iterations of OMP from $y = \Phi x$.

2 Auxiliary lemmas.

We use two results on the rate of convergence of Orthogonal Greedy Algorithm (OMP).

Theorem A. (R.A. Devore, V.N. Temlyakov, [7]) *Suppose that $y = \Phi x$. Then for any $l \ge 1$ we have*

$$\|r^l\| \le |x|_1 l^{-1/2}.$$

Theorem B. (EL, [15]) *For any l, $1 \le l \le 1/(20\mu(\Phi))$ we have*

$$\|r^{2l}\| \le 3\sigma_l(y, \Phi).$$

For $l \ge 0$ we set

$$z^l := x - x^l.$$

Then by definition of OMP

$$r^l = y - \Phi x^l = \Phi x - \Phi x^l = \Phi z^l, \ l \ge 0. \tag{3}$$

Assume that

$$x = (x_1, \dots, x_N), \quad z^l = (z_1^l, \dots, z_N^l), \ l \ge 0.$$

Set

$$V_0 = \mathrm{supp} x, \quad \sharp V_0 \le K. \tag{4}$$

By $x|_V$, $V \subset V_0$, denote an element $(\widetilde{x}_1, \dots, \widetilde{x}_n)$ of $\mathbb{R}^N$ such that $\widetilde{x}_i = x_i$, $i \in V$, and $\widetilde{x}_i = 0$, $i \notin V$. For each $V \subset V_0$ we define

$$R(V) = \sum_{i \in V} x_i^2.$$

Let

$$C := 2 \times 10^5, \quad c := 10^{-6}, \quad \delta := cK^{-0.2}. \tag{5}$$

Lemma 1. *Suppose that $l + K \le CK^{1.2}$. Then we have*

$$\sum_{i \in \Lambda^l} (z_i^l)^2 \le 3\delta R(V_0 \setminus \Lambda^l), \tag{6}$$

$$R(V_0 \setminus \Lambda^l) \le (1 + 2\delta)\|r^l\|^2. \tag{7}$$

Proof. It's clear that $|z^l|_0 \le |x|_0 + |x^l|_0 \le K + l \le CK^{1.2}$, so by RIP and (3) we have

$$(1-\delta)\sum_{i=1}^{N}(z_i^l)^2 \le \|\Phi z^l\|^2 = \|r^l\|^2 \le (1+\delta)\sum_{i=1}^{N}(z_i^l)^2. \tag{8}$$

On the other hand, using definition of $R(\cdot)$, and RIP for $x|_{V_0\setminus\Lambda^l}$ we write

$$\|x|_{V_0\setminus\Lambda^l}\|^2 = R(V_0\setminus\Lambda^l),$$

$$(1-\delta)R(V_0\setminus\Lambda^l) \le \left\|\Phi\left(x|_{V_0\setminus\Lambda^l}\right)\right\|^2 \le (1+\delta)R(V_0\setminus\Lambda^l). \tag{9}$$

The definition of OMP implies that

$$\|\Phi z^l\|^2 = \|r^l\|^2 \le \left\|\Phi\left(x|_{V_0\setminus\Lambda^l}\right)\right\|^2.$$

Therefore using (8) and (9) we have

$$(1-\delta)\sum_{i=1}^{N}(z_i^l)^2 \le \|r^l\|^2 \le \left\|\Phi\left(x|_{V_0\setminus\Lambda^l}\right)\right\|^2 \le (1+\delta)R(V_0\setminus\Lambda^l),$$

$$(1-\delta)\left(\sum_{i\in\Lambda^l}(z_i^l)^2 + \sum_{i\in V_0\setminus\Lambda^l}(z_i^l)^2\right) \le (1+\delta)R(V_0\setminus\Lambda^l).$$

$$\sum_{i\in\Lambda^l}(z_i^l)^2 + R(V_0\setminus\Lambda^l) \le \frac{1+\delta}{1-\delta}R(V_0\setminus\Lambda^l).$$

$$\sum_{i\in\Lambda^l}(z_i^l)^2 \le (\frac{1+\delta}{1-\delta} - 1)R(V_0\setminus\Lambda^l) \le 3\delta R(V_0\setminus\Lambda^l).$$

This completes the proof of (6). From (8) it follows that

$$R(V_0\setminus\Lambda^l) = \sum_{i\in V_0\setminus\Lambda^l}(z_i^l)^2 \le \sum_{i=1}^{N}(z_i^l)^2 \le (1-\delta)^{-1}\|r^l\|^2 \le (1+2\delta)\|r^l\|^2.$$

□

For increasing sequence $0 = l_0 < l_1 < \cdots < l_s$, $s \ge 1$, we denote

$$V_k := V_0\setminus\Lambda^{l_k},\ R_k = R(V_k),\ 0\le k\le s. \tag{10}$$

Lemma 2. *Suppose that $l_k + K \le CK^{1.2}$, $1\le k\le s$. Then for arbitrary $p\in\mathbb{N}$ we have*

$$\|r^{l_k+p}\|^2 \le \frac{R_k}{p}\left(6\delta CK^{1.2} + 2K\right).$$

Proof. Since $r^{l_k} = \Phi z^{l_k}$ we estimate by Theorem A.

$$\|r^{l_k+p}\|^2 \le \frac{|z^{l_k}|_1^2}{p}. \tag{11}$$

So to prove the lemma it's sufficient to estimate

$$|z^{l_k}|_1^2 = \left(\sum_{i=1}^{N} |z_i^{l_k}|\right)^2 = \left(\sum_{i\in V_0\cup\Lambda^{l_k}} |z_i^{l_k}|\right)^2 \le 2\left(\left(\sum_{i\in V_0\setminus\Lambda^{l_k}} |z_i^{l_k}|\right)^2 + \left(\sum_{i\in\Lambda^{l_k}} |z_i^{l_k}|\right)^2\right).$$

Applying (10) and (4) we have

$$\left(\sum_{i\in V_0\setminus\Lambda^{l_k}} |z_i^{l_k}|\right)^2 = \left(\sum_{i\in V_0\setminus\Lambda^{l_k}} |x_i|\right)^2 = \left(\sum_{i\in V_k} |x_i|\right)^2 \le \sharp V_k \sum_{i\in V_k} |x_i|^2 \le \sharp V_0 R_k \le R_k K. \tag{12}$$

Using (6) from Lemma 1 we get

$$\left(\sum_{i\in\Lambda^{l_k}} |z_i^{l_k}|\right)^2 \le \sharp\Lambda^{l_k} \sum_{i\in\Lambda^{l_k}} (z_i^{l_k})^2 = l_k \sum_{i\in\Lambda^{l_k}} (z_i^{l_k})^2 \le CK^{1.2} 3\delta R(V_0\setminus\Lambda^{l_k}) \le CK^{1.2} 3\delta R_k.$$

Combining with (12) we obtain the desirable inequality

$$|z^{l_k}|_1^2 \le R_k(6\delta C K^{1.2} + 2K).$$

This together with (11) completes the proof of the lemma. $\square$

Lemma 3. *Let $1 \le p \le K^{0.8}$ and $l_k + 2p \le CK^{1.2}$, $1 \le k \le s$. Then for any $W \subset V_k$ such that $\sharp W = p$ we have*

$$R(V_k\setminus\Lambda^{l_k+2p}) \le 10R(V_k\setminus W) + 30\delta R_k.$$

Proof. According to RIP, (3), (6) and (10) we estimate

$$\begin{aligned}
\left(\sigma_p(r^{l_k})\right)^2 &\le \|r^{l_k} - \Phi(x|_W)\|^2 = \|\Phi(z^{l_k}) - \Phi(z^{l_k}|_W)\|^2 = \\
&= \|\Phi(z^{l_k} - z^{l_k}|_W)\|^2 \le (1+\delta) \sum_{1\le i\le N,\ i\notin W} (z_i^{l_k})^2 \le \\
&\le (1+\delta)\left(\sum_{i\in V_k\setminus W} (z_i^{l_k})^2 + \sum_{1\le i\le N,\ i\notin V_k} (z_i^{l_k})^2\right) \le (1+\delta)\left(\sum_{i\in V_k\setminus W} (x_i)^2 + \sum_{i\in\Lambda^{l_k}} (z_i^{l_k})^2\right) \le \\
&\le (1+\delta)(R(V_k\setminus W) + 3\delta R(V_0\setminus\Lambda^{l_k})) \le (1+\delta)(R(V_k\setminus W) + 3\delta R_k).
\end{aligned}$$

Since

$$p \le K^{0.8} \le 1/(20\mu(\Phi))$$

we can apply Theorem B and get

$$\|r^{l^k+2p}\| \leq 3\sigma_p(r^{l_k}).$$

Using (7) from Lemma 1 we obtain

$$R(V_k \setminus \Lambda^{l_k+2p}) = R(V_0 \setminus \Lambda^{l_k+2p}) \leq (1+2\delta)\|r^{l_k+2p}\|^2 \leq (1+2\delta)9\left(\sigma_p(r^{l_k})\right)^2 \leq$$
$$\leq (1+2\delta)9(1+\delta)(R(V_k \setminus W) + 3\delta R_k) \leq 10\left(R(V_k \setminus W) + 3\delta R_k\right).$$

□

3 Proof of Theorem 1.

We prove by induction on k that if $l_k + K \leq CK^{1.2}$ and $V_k \neq \emptyset$, then we can define $l_{k+1} > l_k$, V_{k+1} and R_{k+1} satisfying (10) such that

$$l_{k+1} + K \leq CK^{1.2}, \tag{13}$$

and at least one of the following statements hold

$$\text{(A)} \quad l_{k+1} - l_k \leq 3 \times 10^4(6\delta CK^{1.2} + 2K), \quad \sharp(V_k \setminus V_{k+1}) \geq K^{0.8}; \tag{14}$$

$$\text{(B)} \quad l_{k+1} - l_k \leq 2K^{0.8}, \quad \sharp(V_k \setminus V_{k+1}) \geq K^{0.6}; \tag{15}$$

$$\text{(C)} \quad l_{k+1} - l_k \leq 2K^{0.6}, \quad \sharp(V_k \setminus V_{k+1}) \geq K^{0.4}; \tag{16}$$

$$\text{(D)} \quad l_{k+1} - l_k \leq 2K^{0.4}, \quad \sharp(V_k \setminus V_{k+1}) \geq K^{0.2}; \tag{17}$$

$$\text{(E)} \quad l_{k+1} - l_k \leq 2K^{0.2}, \quad \sharp(V_k \setminus V_{k+1}) \geq 1. \tag{18}$$

Set

$$p_A = [3 \times 10^4(6\delta CK^{1.2} + 2K)], \quad W_A = V_k \cap \Lambda^{l_k+p_A}.$$

If $\sharp W_A \geq K^{0.8}$ we define

$$l_{k+1} := l_k + p_A, \quad V_{k+1} := V_0 \setminus \Lambda^{l_{k+1}} = V_k \setminus W_A, \quad R_{k+1} := R(V_{k+1}).$$

Then $\sharp(V_k \setminus V_{k+1}) = \sharp W_A$ and statement (14) holds. The inequality

$$l_k + p_A + K \leq CK^{1.2}$$

will be checked below. Assume that

$$\sharp W_A < K^{0.8} \tag{19}$$

Applying Lemma 2 we have

$$\|r^{l_k+p_A}\|^2 \leq \frac{R_k(6\delta CK^{1.2} + 2K)}{[3 \times 10^4(6\delta CK^{1.2} + 2K)]} \leq \frac{R_k}{2.5 \times 10^4}.$$

Using (7) from Lemma 1 we estimate

$$R(V_k \setminus W_A) = R(V_0 \setminus \Lambda^{l_k+p_A}) \le (1+2\delta)\|r^{l_k+p_A}\|^2 \le (1+2\delta)\frac{R_k}{2.5\times 10^4} \le \frac{R_k}{2\times 10^4}. \quad (20)$$

Set

$$p_B := \sharp W_A, \quad W_B = V_k \cap \Lambda^{l_k+2p_B}.$$

If $\sharp W_B \ge K^{0.6}$ we set

$$l_{k+1} = l_k + 2p_B, \quad V_{k+1} = V_k \setminus \Lambda^{l_{k+1}} = V_k \setminus W_B, \quad R_{k+1} = R(V_{k+1}).$$

Then $\sharp(V_k \setminus V_{k+1}) = \sharp W_B \ge K^{0.6}$ and taking into account (19) we obtain (15). The inequality

$$l_k + 2p_B + K \le CK^{1.2}$$

will be checked below.

Assume that

$$\sharp W_B < K^{0.6}. \quad (21)$$

Applying Lemma 3 for $W = W_A$ and $p = p_B$, and inequality (20) we get

$$R(V_k \setminus W_B) = R(V_k \setminus \Lambda^{l_k+2p_B}) \le 10R(V_k \setminus W_A) + 30\delta R_k \le \frac{R_k}{2\times 10^3} + 30\delta R_k. \quad (22)$$

We repeat these calculations three more times.

Set

$$p_C := \sharp W_B, \quad W_C = V_k \cap \Lambda^{l_k+2p_C}.$$

If $\sharp W_C \ge K^{0.4}$ we set

$$l_{k+1} = l_k + 2p_C, \quad V_{k+1} = V_k \setminus \Lambda^{l_{k+1}} = V_k \setminus W_C, \quad R_{k+1} = R(V_{k+1}).$$

Then $\sharp(V_k \setminus V_{k+1}) = \sharp W_C \ge K^{0.4}$ and taking into account (21) we obtain (16). The inequality

$$l_k + 2p_C + K \le CK^{1.2}$$

will be checked below.

Assume that

$$\sharp W_C < K^{0.4}. \quad (23)$$

Applying Lemma 3 for $W = W_B$ and $p = p_C$, and inequality (22) we get

$$R(V_k \setminus W_C) = R(V_k \setminus \Lambda^{l_k+2p_C}) \le 10R(V_k \setminus W_B) + 30\delta R_k \le \frac{R_k}{2\times 10^2} + 3(10^2+10)\delta R_k. \quad (24)$$

Set

$$p_D := \sharp W_C, \quad W_D = V_k \cap \Lambda^{l_k+2p_D}.$$

If $\sharp W_D \ge K^{0.2}$ we set

$$l_{k+1} = l_k + 2p_D, \quad V_{k+1} = V_k \setminus \Lambda^{l_{k+1}} = V_k \setminus W_D, \quad R_{k+1} = R(V_{k+1}).$$

Then $\sharp(V_k \setminus V_{k+1}) = \sharp W_D \geq K^{0.2}$ and taking into account (23) we obtain (17). The inequality

$$l_k + 2p_D + K \leq CK^{1.2}$$

will be checked below.

Assume that

$$\sharp W_D < K^{0.2}. \tag{25}$$

Applying Lemma 3 for $W = W_C$ and $p = p_D$, and inequality (24) we get

$$R(V_k \setminus W_D) = R(V_k \setminus \Lambda^{l_k+2p_D}) \leq 10R(V_k \setminus W_C) + 30\delta R_k \leq \frac{R_k}{20} + 3(10^3 + 10^2 + 10)\delta R_k. \tag{26}$$

Set

$$p_E := \sharp W_D, \quad W_E = V_k \cap \Lambda^{l_k+2p_E},$$
$$l_{k+1} = l_k + 2p_E, \quad V_{k+1} = V_k \setminus \Lambda^{l_{k+1}} = V_k \setminus W_E, \quad R_{k+1} = R(V_{k+1}).$$

Taking into account (25) we get

$$l_{k+1} - l_k \leq 2K^{0.2}.$$

The inequality

$$l_k + 2p_E + K \leq CK^{1.2}$$

will be checked below.

Applying Lemma 3 for $W = W_D$ and $p = p_E$, and inequalities (26) and (5) we have

$$R_{k+1} = R(V_k \setminus W_E) = R(V_k \setminus \Lambda^{l_k+2p_E}) \leq 10R(V_k \setminus W_D) + 30\delta R_k \leq$$
$$\leq \frac{R_k}{2} + 3(10^4 + 10^3 + 10^2 + 10)\delta R_k < R_k(1/2 + 4 \times 10^4\delta) < R_k.$$

Therefore $W_E \neq \emptyset$ and $\sharp(V_k \setminus V_{k+1}) \geq 1$ and statement (18) holds.

Thus to complete the proof of induction assumption it remains to estimate

$$l_k + p_A + K,\ l_k + 2p_B + K,\ l_k + 2p_B + K,\ l_k + 2p_C + K,\ l_k + 2p_D + K.$$

It's clear that the biggest of these numbers is the first one.

Using induction assumption, inclusion $V_0 \supset V_1 \supset \cdots \supset V_k$, and the equality $\sharp V_0 = K$ we claim that statement (14) (case A) can be fulfilled not more than $K^{0.2} = K/K^{0.8}$ times, statement (15) (case B) not more than $K^{0.4} = K/K^{0.6}$ times, statement (16) (case C) not more than $K^{0.6} = K/K^{0.4}$ times, statement (17) (case D) not more than $K^{0.8} = K/K^{0.2}$ times, and statement (16) (case E) not more than K times. Hence, using (5), we obtain

$$l_k \leq 3 \times 10^4(6\delta CK^{1.2} + 2K)K^{0.2} + 2K^{0.8}K^{0.4} + 2K^{0.6}K^{0.6} + 2K^{0.4}K^{0.8} + 2K^{0.2}K =$$
$$= (18 \times 10^4\delta CK^{0.2} + 6 \times 10^4 + 8)K^{1.2},$$

$$l_k + p_A + K \leq (18 \times 10^4\delta CK^{0.2} + 6 \times 10^4 + 8)K^{1.2} + 3 \times 10^4(6\delta CK^{1.2} + 2K) + K \leq$$
$$\leq (36 \times 10^4\delta CK^{0.2} + 12 \times 10^4 + 9)K^{1.2} \leq$$
$$\leq (36 \times 10^4 \times 10^{-6}K^{-0.2}2 \times 10^5K^{0.2} + 12 \times 10^4 + 9)K^{1.2} < 20 \times 10^4K^{1.2} = CK^{1.2}.$$

Since $\sharp(V_k \setminus V_{k+1}) \geq 1$ there exists $s \in \mathbb{N}$ such that $V_s = \emptyset$ and $l_s + K \leq CK^{1.2}$. Therefore by (10) and (3) we have

$$V_0 \subset \Lambda^{l_s}$$

and

$$r^{l_s} = \Phi z^{l_s} = \Phi(x - x^{l_s}) = 0.$$

Using RIP we finally obtain that

$$x = x^{l_s}. \quad \square$$

Remark 1. We guess that constant 3 in Theorem B is not optimal and hence constants C and c^{-1} from (5) can be reduced.

The author thanks professor V.N. Temlyakov and professor S.V. Konyagin for useful discussions.

References

[1] Baraniuk R. "Compressive sensing". //IEEE Signal Processing Magazine. 2007. V. 24:4. P. 118-121

[2] Baraniuk R., Davenport M., DeVore R.A., Wakin M. "A simple proof of the restricted isometry property for random matrices." //*Const. Approx.* V. 28:3. P. 253263. 2008.

[3] Candès E. "Compressive sampling" //*Int. Congress of Mathematics.* 2006. 3, P. 1433-1452

[4] Candès E., Tao T. "Decoding by linear programming" //*IEEE Trans. Info. Theory.* 2005. V. 51:12, P. 42034215.

[5] Dai W., Milenkovic O. "Subspace pursuit for compressive sensing: Closing the gap between performance and complexity." //*IEEE Trans. Inform. Theory.* V. 55:5. P. 2230-2249. 2009.

[6] Davenport M., Wakin M. "Analysis of Orthogonal Matching Pursuit using the Restricted Isometry Property". //2010. http://arxiv.org/abs/0909.0083

[7] DeVore R. A., Temlyakov V. N. "Some remarks on Greedy Algorithms" // *Advances in Computational Mathematics.* 1996. V. 5. P. 173–187.

[8] Donoho D. "Compressed sensing". //*IEEE Trans. on Information Theory* 2006. V. 52:4. P. 1289 - 1306.

[9] Donoho D. L., Elad M., Temlyakov V. N. " Stable recovey of sparse overcomplete representations in the presense of noise" //*IEEE Trans. Inform. Th.* 2006. V. 52:1. P. 6–18.

[10] Donoho D. L., Elad M., Temlyakov V. N. " On Lebesgue-type inequalities for greedy approximation " //*Journal of Approximation theory* 2007. V. 147:2. P. 185–195.

[11] Gilbert A.C., Muthukrishnan M., Strauss J. "Approximation of functions over redundant dictionaries using coherence" //*Proc. 14th Annu. ACM-SIAM Symp. Discrete Algorithms.* 2003. P. 243-252.

[12] Gribonval R., Nielsen M. " On the strong uniqueness of highly sparse expansions from redundant dictionaries" // *Proc. Int Conf. Independent Component Anal. (ICA'04)* 2004.

[13] Kashin B. "The widths of certain finite dimensional sets and classes of smooth functions." //*Izvestia* V. 41. 334351. 1977

[14] Liu E., Temlyakov V.N "Orthogonal Super Greedy Algorithm and Applications in Compressed Sensing" //2010. http://dsp.rice.edu/sites/dsp.rice.edu/files/cs/LiuTemlyakov.pdf

[15] Livshits E.D. "On the optimality of Orthogonal Greedy Algorithm for M-coherent dictionaries " //2010. http://arxiv.org/abs/1003.5349

[16] Rauhut H. "On the impossibility of uniform sparse reconstruction using greedy methods." //*Sampl. Theory in Signal and Image Process.* V. 7:2. P. 197215. 2008.

[17] Temlyakov V.N., Zheltov P. "On performance of greedy algorithms" // 2010. http://dsp.rice.edu/sites/dsp.rice.edu/files/cs/TemlyakovZheltov.pdf

[18] Tropp J. A. "Greed is good: algorithmic results for sparse approximation" //*IEEE Trans. Inform. Th.* 2004. V. 50:10. P. 2231–2242.